

UC Davis

UC Davis Previously Published Works

Title

Genome-wide DNA methylome variation in two genetically distinct chicken lines using MethylC-seq.

Permalink

<https://escholarship.org/uc/item/2634n3q3>

Journal

BMC genomics, 16(1)

ISSN

1471-2164

Authors

Li, Jinxiu
Li, Rujiao
Wang, Ying
et al.

Publication Date

2015-10-01

DOI

10.1186/s12864-015-2098-8

Peer reviewed

RESEARCH ARTICLE

Open Access



Genome-wide DNA methylome variation in two genetically distinct chicken lines using MethylC-seq

Jinxu Li^{1,5†}, Rujiao Li^{2†}, Ying Wang³, Xiaoxiang Hu¹, Yiqiang Zhao¹, Li Li¹, Chungang Feng¹, Xiaorong Gu¹, Fang Liang², Susan J. Lamont⁴, Songnian Hu⁵, Huaijun Zhou^{3,6*} and Ning Li^{1,7,8*}

Abstract

Background: DNA cytosine methylation is an important epigenetic modification that has significant effects on a variety of biological processes in animals. Avian species hold a crucial position in evolutionary history. In this study, we used whole-genome bisulfite sequencing (MethylC-seq) to generate single base methylation profiles of lungs in two genetically distinct and highly inbred chicken lines (Fayoumi and Leghorn) that differ in genetic resistance to multiple pathogens, and we explored the potential regulatory role of DNA methylation associated with immune response differences between the two chicken lines.

Methods: The MethylC-seq was used to generate single base DNA methylation profiles of Fayoumi and Leghorn birds. In addition, transcriptome profiling using RNA-seq from the same chickens and tissues were obtained to interrogate how DNA methylation regulates gene transcription on a genome-wide scale.

Results: The general DNA methylation pattern across different regions of genes was conserved compared to other species except for hyper-methylation of repeat elements, which was not observed in chicken. The methylation level of miRNA and pseudogene promoters was high, which indicates that silencing of these genes may be partially due to promoter hyper-methylation. Interestingly, the promoter regions of more recently evolved genes tended to be more highly methylated, whereas the gene body regions of evolutionarily conserved genes were more highly methylated than those of more recently evolved genes. Immune-related GO (Gene Ontology) terms were significantly enriched from genes within the differentially methylated regions (DMR) between Fayoumi and Leghorn, which implicates DNA methylation as one of the regulatory mechanisms modulating immune response differences between these lines.

Conclusions: This study establishes a single-base resolution DNA methylation profile of chicken lung and suggests a regulatory role of DNA methylation in controlling gene expression and maintaining genome transcription stability. Furthermore, profiling the DNA methylomes of two genetic lines that differ in disease resistance provides a unique opportunity to investigate the potential role of DNA methylation in host disease resistance. Our study provides a foundation for future studies on epigenetic modulation of host immune response to pathogens in chickens.

Keywords: Epigenetics, DNA methylation, MethylC-seq, Immunity, Chickens

* Correspondence: hzhou@ucdavis.edu; ninglcau@cau.edu.cn

†Equal contributors

³Department of Animal Science, University of California, Davis, CA 95616, USA

¹The State Key Laboratory for Agro-biotechnology, China Agricultural University, Beijing 100193, China

Full list of author information is available at the end of the article

Background

DNA methylation is a central epigenetic modification that occurs in most eukaryotic organisms and plays a crucial role in transcriptional regulation. This epigenetic mark is involved in many cellular processes, including embryogenesis, transposon silencing, genomic imprinting, X chromosome inactivation, and tumorigenesis [1–3].

Of the many approaches to profile genome-wide DNA methylation patterns, MethylC-seq is considered the current “gold standard” [4, 5]. Compared with other methods, this approach achieves higher resolution and more precise methylation levels at the single-base resolution. Due to the significant role of DNA methylation on biological processes, recent studies have focused on genome-wide DNA methylation pattern in different eukaryotic genomes. To date, single-base resolution DNA cytosine methylome maps for *Arabidopsis*, human, silkworm, and chicken have been generated by MethylC-seq [3, 6–9].

The chicken (*Gallus gallus*), as a representative of extant avian species, is a model organism [10] for studying embryology, immunology, behavior, and reproduction [11, 12]. Furthermore, avian species are reservoirs of many zoonotic pathogens and, therefore, studies of their genomes provide new insights into how the host responds to pathogens of biomedical importance [11, 13, 14]. Although the chicken genome has been sequenced, one-dimensional sequencing information can only address a fraction of the relevant biological questions [15]. The chicken genome has an active DNA methylation system [16, 17]; therefore, it is of interest to decipher the methylation-determined epigenetic landscape of the chicken. Whole-genome DNA methylation profiling of the chicken was previously conducted using MeDIP and Methyl-MAPS [18–20]. Because of the technical limitations inherent in these techniques, such as low resolution and/or complexity, only a reduced landscape of the DNA methylome was

generated. Recently, single-base resolution DNA methylome sequencing of chicken sperm cells was accomplished [9]. Because of the unique nature of high methylation levels in sperm cells, however, single-base high-resolution methylome information on other tissues in the chicken is needed to enhance our understanding of genome-wide profiles of DNA methylation in this species.

Disease resistance has been one of most challenging traits to enhance in the poultry breeding industry. During domestication and genetic selection, different chicken breeds have displayed a variety of resistance levels to pathogens. For example, the Fayoumi chicken, which originated in Egypt, has been demonstrated to resist viral, bacterial and parasitic infections including Marek's disease virus, avian influenza virus (AIV), *Salmonella enteritidis* and *Eimeria* coccidiosis [21–25]. The Leghorn line used in the current study, which was derived from egg-laying stock in the U.S., is relatively susceptible to pathogen infection compared to the Fayoumi [26].

Previous studies have suggested that abnormal DNA methylation contributes to cancer and infectious disease [27, 28]. To explore potential regulatory roles of DNA methylation related to disease resistance, we generated whole genome single-base DNA methylation profiles of two highly inbred Fayoumi and Leghorn lines (inbreeding coefficient >99.99 %). These two chicken lines have been studied for many facets of disease resistance and immune response [26, 29–31]. Their high inbreeding level minimizes within-line genetic variation and their distinct pathogen responses between lines provide an opportunity to identify DNA methylation differences between disease resistant and susceptible chicken lines and to explore the potential biological role of DNA methylation on immune response. In addition, transcription profiling of the same tissues used in the DNA methylation analysis was

Table 1 Sequencing results and read alignment

Samples	Library	Reads	Reads after filtered	Mapped reads	Uniq mapped reads	Post-processed	Bisulfite conversion rates	CpG coverage %
Fayoumi replicate 1	A	175421846	159533236	144001914	114988172	101197098	99.21 %	83.72 %
	B	175900866	151260322	118878023	79759740	68021164		
Fayoumi replicate 2	A	172418670	153201632	135605330	108313076	94229642	99.24 %	85.01 %
	B	170400106	151853856	127905703	96985346	85265402		
	All	694141488	615849046	526390970	400046334	348713306		
Leghorn replicate 1	A	200045434	178899768	160144382	126118192	87060156	99.33 %	86.92 %
	B	156119182	136244806	122471089	98190500	88847236		
Leghorn replicate 2	A	516069410	448675100	299761294	126087800	37391344	99.40 %	91.57 %
	B	293866348	268269302	239807905	191655228	151160202		
	All	1166100374	1032088976	822184670	542051720	364458938		

The chicken reference genome was downloaded from UCSC database (<http://genome.ucsc.edu>) Nov. 2011 (ICGSC *Gallus_gallus*-4.0/galGal4)

performed to interrogate the relationship between DNA methylation and transcriptional regulation on a genome-wide scale.

Results

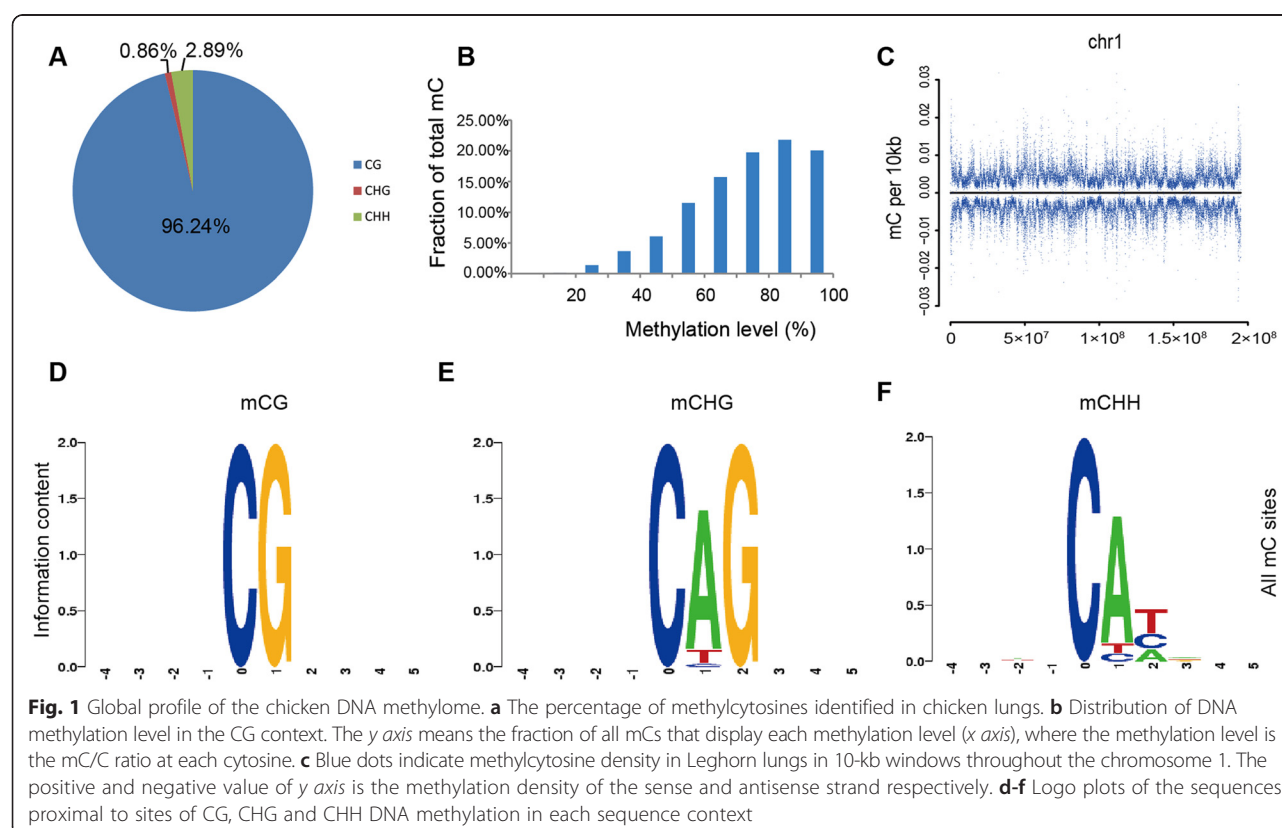
Single-base resolution DNA methylome of chicken lungs

Because the lung is one of major tissues where avian influenza virus (AIV) replicates in the chicken, we characterized whole-genome single-base DNA methylation profiles of chicken lung tissues from Fayoumi and Leghorn by the MethylC-seq. The same tissue samples were used in a previous study [25]).

A total of 148.82 gigabases (Gb) of sequence were generated from two biological replicates of each of two genetic lines. After filtering, 0.40 and 0.54 billion reads from Fayoumi and Leghorn lines, respectively, were uniquely mapped to the galGal4 reference sequence. Reads including more than three cytosines in non-CG contexts were considered as non-converted and removed. Then, we obtained 0.35 and 0.36 billion reads with an average read depth of 13.9 \times and 14.5 \times per strand for Fayoumi and Leghorn, respectively (Table 1). The bisulfite conversion rates for all samples were 99.21 to 99.4 %. We used a binomial distribution to identify the methylcytosines, and 1 % of false discovery rate (FDR) was used to correct it. To validate the accuracy and repeatability of the methylation

profiles, we compared the mCs identified independently in the two biological replicates, and found that nearly 90 % mCG sites were identical in the two individuals within each line, but only half of the mCs at non-CG sites were shared between lines (Additional file 1). To ensure the accuracy of our results, the intersection of methylcytosines in replicate 1 and replicate 2 were defined as mCs in each line. Finally, we detected approximately 11.3 and 11.7 million methylcytosines in Fayoumi and Leghorn, respectively, which represented about 3 % of all cytosines obtained by sequencing. More than half of the cytosines in CG contexts were methylated, whereas the methylation rates of the cytosines in CHG and CHH contexts (where H is A, C, or T) were only 0.11 and 0.12 %, respectively (Additional file 2). A total of 96.24 % of all methylcytosines occurred in the CG context, 0.86 % in the CHG context, and 2.89 % in the CHH context (Fig. 1a). To validate MethylC-seq results, we randomly examined methylation level of 165 mCGs, 61 mCHGs and 42 mCHGs using bisulphite PCR sequencing (BS-PCR). Of all the validated 165 mCGs, 96 % confirmed the sequencing results. However, none of the non-CG mCs were validated by the BS-PCR (Additional file 3).

We defined the proportion of reads covering each methylcytosine relative to the total number of reads covering the sites as the methylation level given a specific



cytosine. In the chicken lung, 62 % of mCG sites were 70–100 % methylated (Fig. 1b). We used chromosome 1 as an example to demonstrate the chromosome-wide DNA methylation density, and the CG methylation level revealed large variations across the chromosome 1, which was the same as other chromosomes (Fig. 1c, Additional file 4). The strand-specific mCGs were analyzed, where two strands showed a symmetrical methylation pattern. We also analyzed the genome sequence preference proximal to the sites of methylated CG and non-CG contexts. No sequence preference was found in the mCG-flanking regions or upstream of non-CG methylation; however, the base following a non-CG methylcytosine was almost always an adenine, while thymine was observed less often (Fig. 1d–f).

Gene methylation profile

Methylation of CpG islands plays an important role in gene regulation during development, and it is altered in disease states [2, 32–34]. The CpG islands have been extensively studied since their identification more than 20 years ago [35]. In the present study, we characterized the CpG island methylation pattern in the chicken genome. To be characterized as a CpG island, a sequence must meet the following criteria: 1) (G + C) content above 55 %; 2) observed CpG/expected CpG of 65 % or greater; 3) more than 500 bp in length (<http://methycancer.psych.ac.cn/CpG130.do>) [36]; and 4) for methylated CpG island, the methylated CpG/ total CpG dinucleotide of 70 % or greater. By this definition, we identified 3767 methylated CpG islands. Among the 24,222 CpG islands, more than 80 % were un-methylated. Of all the methylated CpG islands, most (69.74 %) were in the intergenic regions. Only 5.04 % were in the gene upstream region and 3.90 % were in the downstream region, indicating that CpG islands at the 5' and 3' ends of genes were generally un-methylated. CpG islands in the gene-body

regions (21.32 %) were more methylated than those in the 5' and 3' UTR regions (Fig. 2a).

To characterize methylation of chicken genes, we calculated the relative methylation levels (mC/CG) in the context of gene regions and of their upstream and downstream regions. We further divided the gene-body region into the first exon, first intron, internal exons, internal introns, and last exon. In general, the relative methylation level was higher in the gene-body regions than in the 5' upstream and 3' downstream regions. For the gene body, the relative methylation level was relatively low in the first exon and higher in the first intron. It reached the highest level in the internal exons and remained at a high level until the transcription termination site. Interestingly, there was always a sharp decrease of methylation across the exon-intron boundaries (dashed lines 2–5) (Fig. 2b).

We next investigated the promoter methylation levels of different gene categories. A promoter region was defined as −1.5 kb to 0.5 kb relative to the TSS. Compared with protein-encoding genes, microRNA (miRNA) ($P < 2.2E^{-16}$) and small nucleolar (snoRNA) ($P = 1.99E^{-14}$) were more highly methylated, and the tRNA ($P = 9.56E^{-4}$) genes were less methylated, whereas miscellaneous RNA (misc_RNA) ($P = 0.84$), rRNA ($P = 0.15$) and snRNA ($P = 0.08$) did not show significant variance ($P < 0.05$) (Fig. 3). Our results confirmed that miRNA genes were typically highly methylated [8, 37], suggesting that this phenomenon is highly conserved among species.

DNA methylation and gene evolution in the chicken genome

To characterize evolutionary changes in gene methylation, we further classified chicken genes into four temporal groups based on a nucleotide sequence similarity search using BLAST against several clades in the evolutionary tree (hereafter referred to as TG, with TG1 being the oldest group; see the Methods section) [38]. We used the gene sequence to construct temporal groups. After that,

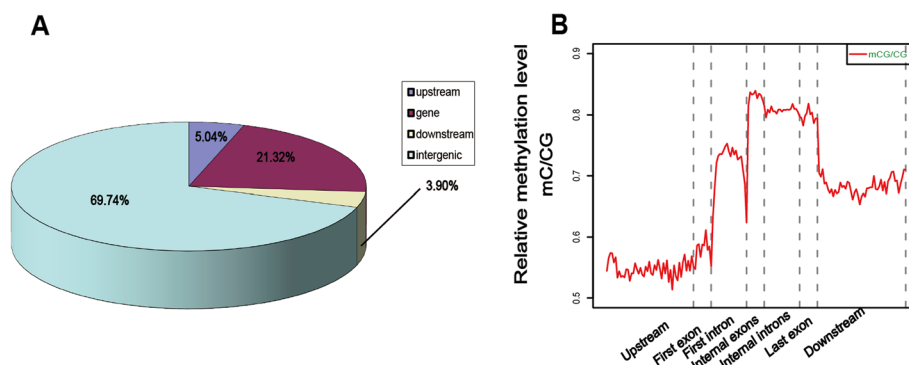


Fig. 2 Distribution of methylated cytosines in different genome regions. **a** Proportion of methylated CpG islands in different genomic regions. **b** Relative methylation level in gene regions (Different areas were divided by dotted lines)

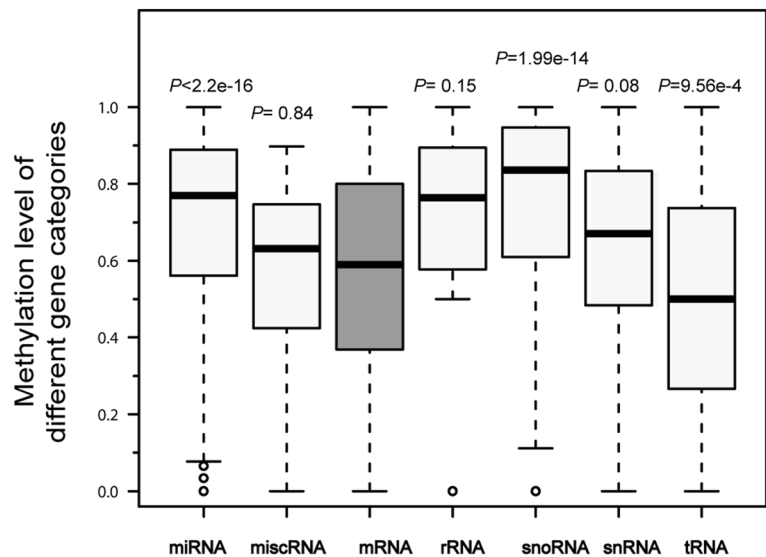


Fig. 3 Promoter relative methylation level of different gene categories in the chicken genome. Box plots showed the methylation level of each gene category. Each category was compared with coding protein. miRNA ($P<2.2E-16$), misc_RNA ($P=0.8381$), rRNA ($P=0.1472$), snoRNA ($P=1.99E-14$), snRNA ($P=0.08$) and tRNA ($P=9.56E-4$)

we investigated the promoter and gene body methylation level of different temporal groups. Promoter methylation of the evolutionarily oldest genes had a significantly lower level (Student's t-test, P -values were showed in Table 2), while the methylation level tended to increase with the evolution of genes (Fig. 4a). Gene body methylation level of different temporal groups was also performed and showed that gene body methylation of newly evolved genes was higher than early evolved groups (Fig. 4b).

The GO annotation analysis was conducted for hyper-methylated genes (methylation level $\geq 70\%$) and hypo-methylated genes (methylation level $\leq 30\%$) using WEGO (Web Gene Ontology Annotation Plot <http://wego.genomics.org.cn/cgi-bin/wego/index.pl>). Promoter and gene-body regions were separately analyzed. The gene list of each group is shown in Additional file 5. For the promoter region, most of the GO groups have more

hypo-methylated genes, especially in cellular development, biological regulation and metabolic process. However, the molecular function with transducer and receptor activity, which is related to cellular responding to stimuli, establishment of localization, signaling and immune system processes, tend to have more hyper-methylated genes (Additional file 6). For the gene-body regions, almost all genes are hyper-methylated (Additional file 6), while hypo-methylated genes, which are rare, did not show any enrichment.

In addition, we also utilized DAVID functional annotation tool to perform the enrichment analysis (<http://david.abcc.ncifcrf.gov/>). The results were consistent with WEGO analysis, which showed that promoter hyper-methylated genes were significantly enriched in biological processes of stimuli, such as cognition, sensory perception and defense response. In the meanwhile, promoter hypo-methylated genes were clustered in transcription regulator activity (Additional file 7).

Table 2 P -value between every two temporal gene groups

Promoter	TG2	TG3	TG4
TG1	4.40E-08	<2.2e-16	<2.2e-16
TG2		1.30E-08	<2.2e-16
TG3			6.21E-07
TG4			
Gene body	TG2	TG3	TG4
TG1	3.08E-09	2.26E-13	<2.2e-16
TG2		0.0003188	<2.2e-16
TG3			0.01255
TG4			

DNA methylation distribution in the repeat elements and pseudogenes

DNA methylation is essential for silencing transposable elements and other repetitive elements in eukaryotes. To investigate the effect on the regulation of repeat elements caused by DNA methylation in the chicken, we first confirmed the methylation level of repeat elements and its flanking regions. The absolute methylation level (mC/Length) in repetitive elements displayed a lower level than their flanking regions. With regard to the relative methylation level (mC/CG), the methylation

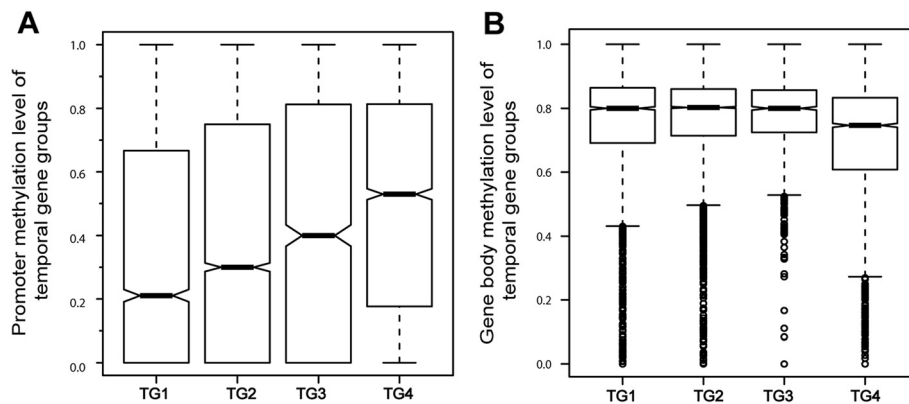


Fig. 4 a Promoter relative methylation level of different temporal groups. The species used for each temporal group were: TG1 (African malaria mosquito, fruitfly, nematode, *Schistosoma* and yellow fever mosquito), TG2 (medaka, pufferfish, trout and zebrafish), TG3 (clawed frog and tropical frog), TG4 (all chicken genes not found in the above species). **b** Gene body relative methylation level of different temporal groups

density was lower in overall repetitive elements regions with the exception of a sharp increase in the boundary regions (Fig. 5a).

We then further compared the relative methylation level of different types of repeat elements (downloaded from UCSC database <http://hgdownload.soe.ucsc.edu/downloads.html#chicken>) and genome-wide randomly selected regions. For randomly selected regions, we excluded repeat sequences and genic regions (from 2 kb upstream of TSS to 2 kb downstream of TTS) to avoid the bias of sequence feature. In addition, the length and the number of repeat elements were also considered in the randomly selected regions. We observed that in chickens, all repeat elements had a lower relative methylation level than the randomly selected regions, except for mariner DNA repeats (Additional file 8).

We next compared the methylation level of pseudogenes with their corresponding genuine genes. The pseudogene information of chicken was downloaded from the pseudogene database (www.pseudogene.org) [39, 40], and the

sequence information of pseudogenes and genuine genes were downloaded from the UCSC database. The pseudogenes' promoter methylation levels were significantly higher than genuine genes ($P = 3.645 \times 10^{-5}$) (Fig. 5b).

Correlation between DNA methylation and gene expression

To analyze the correlation between average methylation degree (average methylation level of each CG sites) and expression of gene at mRNA level in these two genetic lines, RNA-seq profiles of chicken lungs from the same individuals that were used for MethylC-seq were generated. We divided the genes into five groups according to the mRNA expression level, from the bottom 20 % to the top 20 %, corresponding to the 1st to 5th quintiles. In general, the methylation degree across the five groups began to decrease from 1 kb upstream of the transcriptional start site (TSS) of the genes, and it increased after TSS (Figs. 6a, b). Box plotting showed that DNA methylation in the promoter region was negatively correlated

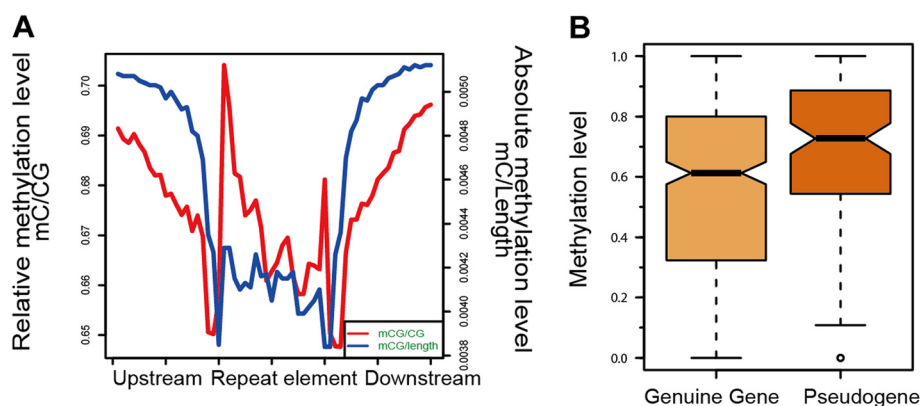
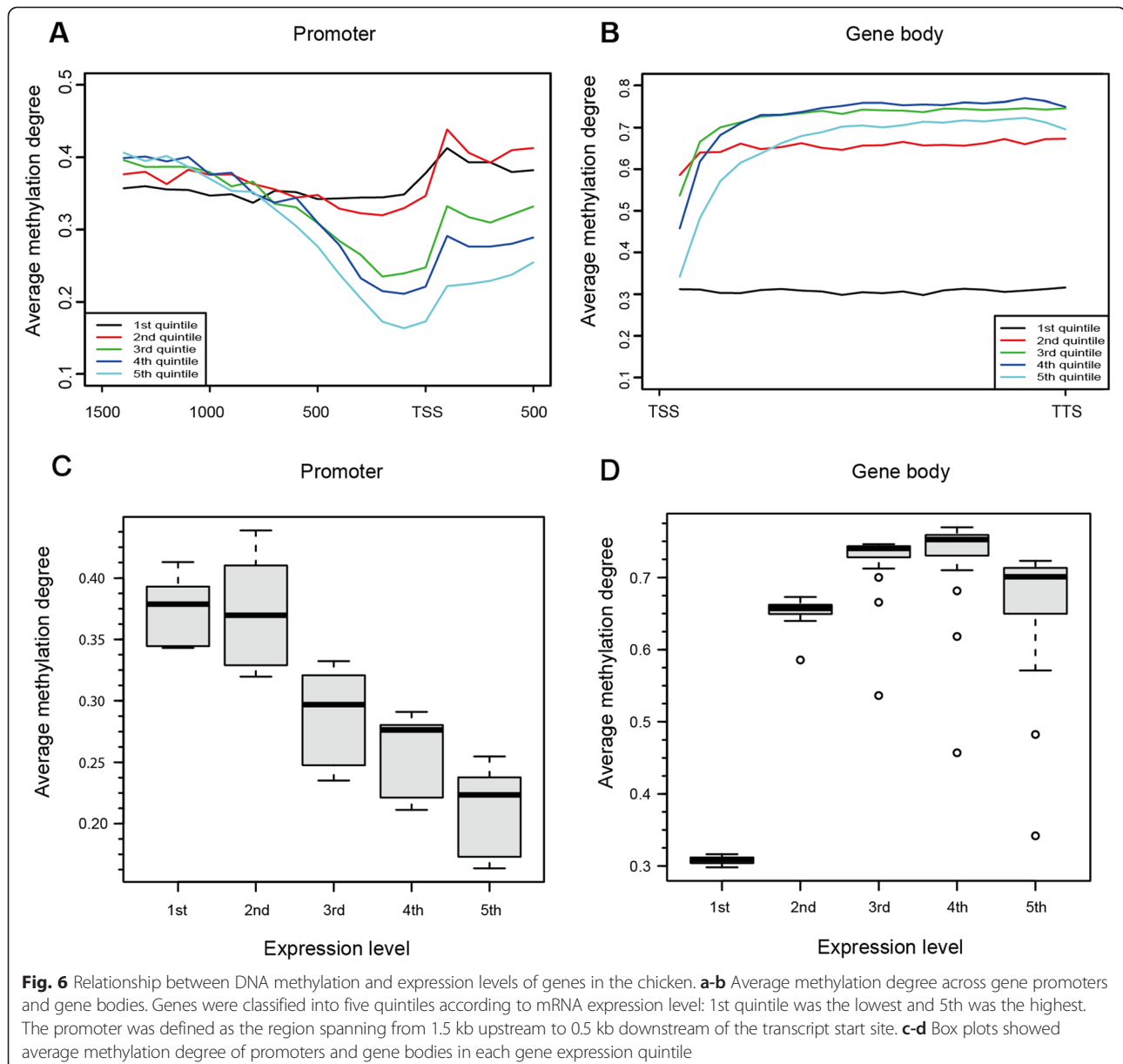


Fig. 5 DNA methylation distribution in the repeat sequences and pseudogenes. **a** Absolute methylation level (blue line) and relative methylation level (red line) in repeat elements regions. **b** Promoter relative methylation level of pseudogenes and corresponding genuine genes



with mRNA expression, especially at 500 bp around TSS (Figs. 6a, c). In contrast, the correlation between gene-body methylation and mRNA expression was more complex. In general, the expression level of the moderately expressed groups (2nd, 3rd and 4th quintiles) was positively correlated with the gene methylation. However, the methylation degree in the 5th quintile (highest expressed group) was lower than the 3rd and 4th quintiles (Fig. 6b, d). Methylation of the 1st quintile was much lower than other groups.

DNA methylation differences between two genetic lines

We characterized the methylation differences between the two genetic lines and explored how these methylation

differences affected gene expression differences. To identify DNA methylation differences between the two genetic lines, we applied a sliding window method to identify the differentially methylated regions (DMRs) between Fayoumi and Leghorn chickens (5 % FDR). One-kilobase windows that contained at least four differential mCGs were identified, and adjacent windows were merged (see the Methods section). A total of 5652 DMRs were identified, among which 2400 DMRs were located within RefSeq genes (from 2 kb upstream to 2 kb downstream of gene). Based on these DMRs, we obtained 1532 DMR-associated genes (Additional file 9). Combined with RNA-seq results of the two lines [25], the DMRs were associated with 705 and 744 genes more highly expressed in Fayoumi and Leghorn,

Table 3 GO enrichment of DMR-associated genes

Gene ontology (GO) term	P-value for enrichment
Neuron differentiation	3.60E-04
Immunoglobulin domain	6.80E-04
Cell morphogenesis involved in differentiation	7.20E-04
Phosphate metabolic process	8.00E-04
Immune effector process	1.20E-03
Nucleoside binding	1.40E-03
Leukocyte mediated immunity	1.60E-03
Cellular component morphogenesis	2.10E-03
Enzyme activator activity	2.40E-03
Organelle lumen	3.00E-03

Promoter and gene body methylation level were calculated separately

respectively (of 1532 DMR-associated genes, 83 showed no expression in both two chicken lines) (Additional file 10). And, 190 differentially methylated genes had more than two-fold differences in mRNA expression (Additional file 11).

We randomly selected *HCK* (a DMR-associated gene) to represent the DNA methylation and gene expression difference between the two chicken lines. MethylC-seq and BS-PCR validation results of ENSGALG00000006522 (*HCK*) were performed (Additional file 12), and differentially methylated region was highlighted by light gray (Additional file 12). We found that in the DMR, the methylation level of Fayoumi was higher than Leghorn and the BS-PCR validation result displayed an even more pronounced difference. In contrast, RNA-seq results showed that the expression of *HCK* was three times higher in Leghorn than in Fayoumi (Additional file 12).

The differentially methylated genes were further analyzed by gene ontology enrichment analysis using *DAVID*. Several immune-related GO terms including immunoglobulin domain, immune effector process and leukocyte mediated immunity were significantly enriched (Table 3). Of particular interest, several immune-related genes such as *TLR 4* and *PIK3CD* were both differentially methylated and differentially expressed between Fayoumi and Leghorn lines (Additional file 13).

To evaluate the genome-wide gene expression and DNA methylation differences between Fayoumi and Leghorn, we obtained *p*-values of each gene between the two lines for transcription and DNA methylation by χ^2 test and generated the distribution diagram. The results showed that the degree of DNA methylation variation between the two lines across the genome was greater than gene expression variation (Fig. 7). Furthermore, we analyzed the correlation between the whole genome promoter DNA methylation differences and gene expression differences but found no significant correlation (Additional file 14).

Discussion

The stability of DNA methylation plays an important role in preventing tumorigenesis and disease progression [27, 28]. These results suggest that mediating gene regulation by DNA methylation may be associated with disease resistance. To improve our understanding of the relationship between DNA methylation and breed-specific disease resistance, we analyzed whole-genome single-base resolution DNA methylomes of Fayoumi and Leghorn chicken lungs. In this study, we report the single base DNA methylomes of chicken lungs and interrogate the potential role of DNA methylation on immune response.

Chicken methylomes have been previously characterized using MeDIP and Methyl-MAPS, which are techniques based on antibody binding affinity and restriction enzyme digestion [18–20]. Although both methods perform well for CpG-rich regions, they generate much lower resolution and coverage. The CpG coverage was only 32 % from the study in chickens using Methyl-MAPS [19]. In the current study, a total of 148.82 Gb sequencing data were generated, and the CpG coverage for each biological replicate ranges from 83.72 to 91.57 % (Table 1), which is comparable to both the human (94 %) and silkworm (92 %) methylomes [3, 8].

In the chicken, the general DNA methylation pattern is consistent with other species. For example, cytosine methylation occurs almost exclusively in the CG contexts; gene-body exhibits higher methylation than the 5' and 3' flanking regions [41–43]; promoter methylation negatively correlates with gene expression. These results suggest that the transcriptional regulatory role of DNA methylation is conserved among species [15, 44]. For the gene body regions, we found that internal exons show the highest methylation level, which is consistent with the study using Methyl-MAPS [19]. In addition, there was a great fluctuation of methylation across the exon-intron boundaries, which implies a potential link between DNA methylation and splicing [45]. In contrast with promoters, gene body methylation is positively correlated with gene expression except for the highest expressed group. It is in agreement with previous studies that both lowly and highly expressed genes have a low level of gene body methylation [5, 41, 42, 46]. A study in *Arabidopsis thaliana* suggests that this phenomenon was associated with formation of pre-initiation complexes and was directed by the siRNA pathway [47].

The results also revealed that the chicken DNA methylome has some different features than other species. Previous studies showed that one of primary functions of DNA methylation is host genome defense and targeting the endogenous transposable elements [48–51]. Further studies have proved that DNA methylation is a key regulator of transposon silencing in plants, some animals, and fungi, but not in invertebrates [42, 52]. In the chicken genome,

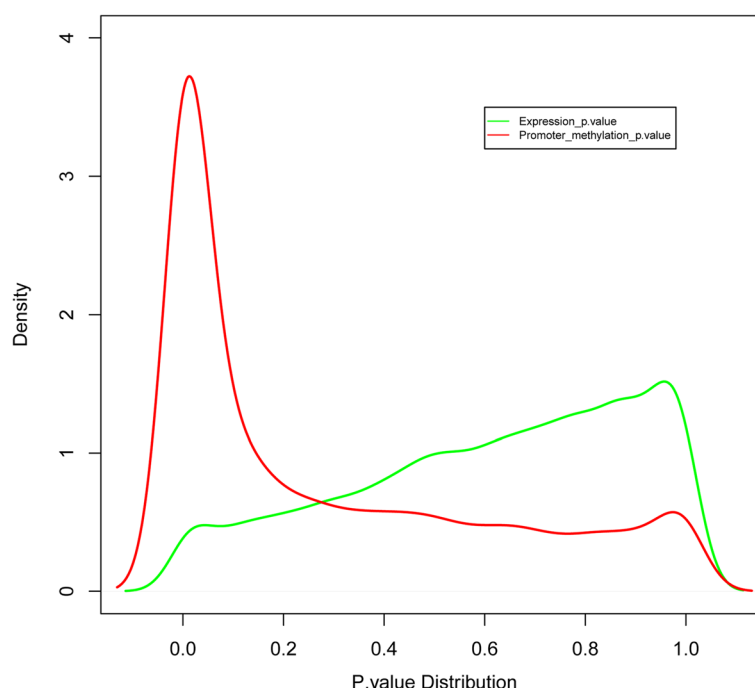


Fig. 7 Genome wide gene expression and DNA methylation variation degree between Fayoumi and Leghorn. The transcription and DNA methylation p -value of each gene between two lines were calculated by χ^2 test

the density of interspersed repeat elements is clearly lower than that in mammalian genomes [11]. Our study showed that the methylation level of repeat elements was lower than both their flanking regions and genome randomly selected regions. These results suggest that the phenomenon of DNA hyper-methylation in repeat elements existing in other vertebrates [42] may not exist in chickens. The chicken genome has relatively low transposable element activity [11]; therefore, we speculate that there may be no need to silence transposable elements through DNA methylation in the chicken.

In this study, we compared the promoter methylation levels in different gene categories. The results confirmed that the promoter regions of miRNAs in the chicken were highly methylated. Furthermore, analysis of the promoter methylation level of genuine genes and their corresponding pseudogenes showed that the methylation of pseudogenes was higher than that of genuine ones. Our results suggest that promoter hyper-methylation of miRNA and pseudogenes may suppress the expression of these genes, which is important to maintain the stability of chicken genome. In addition, we analyzed the methylation level changes related to gene evolution in chickens. The results suggest that the extent of gene methylation level was altered in different evolutionary stages. The newer gene group had higher methylation levels in the promoter regions, while the conserved ones had lower methylation levels. This result may arise from the phenomenon that ancient genes tend to be constitutively expressed [38, 53]

and promoter methylation may be dispensable for that process. In contrast, the newest genes group is more likely to be tissue-specifically expressed [54] and, therefore, they may be more dependent on methylation regulation. Furthermore, we found that the gene body methylation level of the evolutionarily conserved genes was higher than that of the newest set of genes. This agrees with previous studies that demonstrated that gene body hyper-methylated genes were conserved and were functionally important [55, 56].

To expand our knowledge on the potential relationship between DNA methylation and disease resistance in chickens, the differentially methylated regions between two chicken lines that differ for disease resistance were identified. The many immune-related GO terms significantly enriched from DMR-associated genes between the two genetic lines suggest that DNA methylation may serve as one of the regulatory mechanisms that modulates immune response. Of particular note, some of the immune-related genes within DMRs also had significant mRNA expression differences between the two lines, including *PIK3CD* and *TLR4*. *PIK3CD* has been reported to play important roles in both innate and adaptive immunity; *PIK3CD* mutant mice showed decreased immune responses and impaired B/T cell development and function [57, 58]. *TLR4* (Toll-like receptor 4) is a member of the TLR family and plays a major role in pathogen recognition and activation of innate immunity [24]. Moreover, susceptible chickens have been reported to have an increased

methylation level of the *TLR4* gene after *Salmonella* infection [59]. The results collectively suggest that DNA methylation may regulate host immune response via modulating expression of certain immune-related genes.

Of 1532 DMR-associated genes between the two lines, only 190 genes had significant expression differences (\geq two-fold changes). This finding is similar to previous studies in which only 6 % of differentially methylated genes had significant expression differences in *Arabidopsis thaliana* ecotypes, and 7.1 % in cultivated and wild rice [60, 61]. However, no significant correlation existed between genome wide-gene expression differences and promoter DNA methylation differences of the two lines, likely because DNA methylation is only one of many factors regulating transcription. Finally, a large number of genes showed significant differences in DNA methylation between the genetic lines, while a limited number of genes had significant differences in mRNA expression (Fig. 7). This suggests that DNA methylation may contribute more genetic differences between the two lines than transcription.

Conclusion

This study provides the first report of single-base resolution methylation profiles in chicken tissues, which may serve as chicken reference epigenomes. We illustrate the regulatory role of DNA methylation in controlling gene expression and maintaining genome transcription stability. By profiling DNA methylomes of two unique highly inbred lines, our results also suggest the potential role of DNA methylation in regulating disease resistance in chickens.

Methods

Biological materials

Two genetically distinct, highly inbred chicken lines (Leghorn GB2 and Fayoumi M43) with an inbreeding coefficient of more than 99.99 % were used [26]. Two Leghorn and two Fayoumi birds (one male and one female chicken of each line) were euthanized at 3 weeks, and lungs were harvested. The animal experiment was performed according to the guidelines approved by the Institutional Animal Care and Use Committee, Texas A&M University.

DNA preparation and MethylC-seq library generation

DNA from lung tissue was isolated by phenol-chloroform extraction. Five μ g DNA was sonicated to produce 200–400 bp fragments, followed by end repair with a nucleotide triphosphate mix free of dCTP. Cytosine methylated adapters provided by Illumina (Illumina, San Diego, CA) were ligated to the sonicated DNA according to the manufacturer's instructions for genomic DNA library construction. Adapter-ligated DNA with the length of 320–500 bp was isolated by 2 % agarose gel electrophoresis,

and sodium bisulfite conversion was performed using the MethylEasy Xceed kit (Human Genetic Signatures, NSW, Australia) according to the manufacturer's instructions. Then 300 ng of bisulfite-converted, adapter-ligated DNA molecules were enriched by 14 cycles of PCR with the following reaction composition: 2.5 U of uracil-insensitive Pfu TurboCx Hotstart DNA polymerase (Stratagene), 5 μ L of 10 \times Pfu Turbo reaction buffer, 25 μ M dNTPs, 1 μ L of Primer 1.1, and 1 μ L of Primer 2.1 (50 μ L final). The thermocycling parameters were as follows: 98 $^{\circ}$ C for 2 min, followed by 4 cycles of 98 $^{\circ}$ C for 15 s, 60 $^{\circ}$ C for 30 s, and 72 $^{\circ}$ C for 1 min, ending with incubation at 72 $^{\circ}$ C for 10 min. The reaction products were purified using the MinElute PCR purification kit (Qiagen, Valencia, CA) and then separated by 2 % agarose gel electrophoresis and purified by the MinElute gel purification kit (Qiagen, Valencia, CA).

High-throughput sequencing

The DNA libraries were sequenced using the Illumina Genome Analyzer II (GA II) according to the manufacturer's protocols. The MethylC-seq libraries were subjected to 80 or 81 cycles to yield longer sequences that are more amenable for unambiguous mapping to the galGal4 genome reference sequence. The chicken reference genome was downloaded from UCSC database (<http://genome.ucsc.edu>) Nov. 2011 (ICGSC Gallus_gallus-4.0/galGal4). Two independent libraries from each biological replicate (line) were sequenced so that parallel analysis was performed on each biological replicate to insure the accuracy of sequencing result.

Processing and alignment of MethylC-seq

Read sequences produced by the Illumina pipeline in the FastQ format were first pre-processed. Reads containing more than three cytosines in a non-CG context were considered unconverted sequences and removed [3]. Then, the reference sequence was prepared and simultaneously converted twice as follows: (1) cytosines were replaced with thymines and (2) guanines were replaced with adenines. Following pre-processing, the reads were sequentially aligned to two computationally converted reference sequences using the BWA. All results from the alignment of a read to both the Watson and Crick converted genome sequences were combined, and if more than one alignment position existed for a read, it was categorized as ambiguously aligned and disregarded. We removed reads that shared the same 5' alignment position within each library, referred to as "clonal" reads, leaving the first read. Reads mapped to the wrong strands were discarded (T-rich reads mapped to Crick-strand Cs converted to Ts or to Watson-strand Gs converted to 'A's, A-rich reads mapped to Watson-strand Cs converted to Ts or to Crick-strand Gs converted to 'A's) [62]. Subsequently, the reads

from all libraries of a particular sample were combined. All unambiguous, or “unique”, read alignments were then subjected to post-processing.

Identification of methylated cytosines

To identify methylcytosines, we used binomial distribution and then 1 % false discovery rate (FDR) was used to correct the *P*-value. We kept the number of false positives methylcytosine calls below 1 % of the total number of methylcytosines identified. The probability *p* in the binomial distribution $B(n, p)$ was estimated from the number of cytosine bases sequenced in reference cytosine positions in the unmethylated mitochondria genome (Error rate: non-conversion plus sequencing error frequency) [63, 64]. For each reference cytosine, the number (*n*) is the read depth, and the cytosine is noted as methylated if the number of sequenced cytosines (*m*) follows the following formula as below [62]:

$$C_n^m p^m (1-p)^{n-m} < 0.01m/(n-m)$$

For each biological replicate, the reads from two technical replicates (A and B) were pooled to provide greater coverage for the identification of the methylcytosines. The methylcytosines presented in this study represent the consensus between two biological replicates.

Identification of differentially methylated regions (DMRs)

We first used Fisher's Exact Test to find the differentially methylated cytosines between Fayoumi and Leghorn with 5 % FDR correction and at least two-fold difference in methylation level (mCG/CG reads). Then, a sliding window method was used to search for differentially methylated regions between Fayoumi and Leghorn line. A 1 kb window containing at least four differential mCGs was considered as an initial seed, and moved 100 bp per iteration in both the 5' and 3' directions. When a 1 kb window containing at least four differential mCGs was identified, the region was extended in 100 bp increments until a 1 kb increment was reached that contained less than four differential mCGs. After the extension in both directions, regions that contained at least 5 differential mCGs and were at least 1 kb long were identified as DMRs [3]. These DMRs were joined together by removing the overlapping region on the same chromosome.

cDNA Library Preparation and Sequencing by RNA-Seq

Two lung RNA samples from the same genetic line were pooled to generate a total of two pooled RNA samples. Total RNA (7 µg) was subjected to two rounds of hybridization to oligo (dT) beads (Invitrogen, Carlsbad, CA) to enrich mRNA. Ribosomal RNA contamination was evaluated by RNA pico chip using a BioAnalyzer (Agilent, Santa Clara, CA). The resulting mRNA was then used to

prepare cDNA libraries using the RNA sequencing sample preparation kit (Illumina, San Diego, CA). The two libraries were sequenced by Illumina Genome Analyzer II, which generated two datasets.

Data filtering, mapping reads and transcriptome analysis

The sequences generated were initially subjected to a filtering process. Any reads that contained numerous interspersed Ns in their sequences, or had relatively short reads (<17 bp), were removed for subsequent analysis. Sequence reads obtained after quality control with filtering were analyzed using CLC Genomics Workbench 4 (CLC bio, Cambridge, MD). After mapping, the unique gene reads for all of the 17,108 annotated chicken genes in the database from the two libraries were combined, and analyzed using the DESeq R package [65]. Differentially expressed genes between two genetic lines were identified at combined fold change >2. The RNA-seq number showed in Additional file 10 is normalized unique gene reads. Genes with low absolute reads (below 10) had been removed. Statistics related to over representation of functional categories were performed using DAVID [66–68].

Chicken gene temporal groups construction

All chicken genes were divided into four temporal groups based on the nucleotide sequence similarity according to the clades in the evolutionary tree (including insect, fish, amphibians, birds and mammals) and searched using BLAST with an E-value threshold set to e^{-20} . Specifically, if a chicken gene had sequence homology with fruitfly, mosquito, and nematode or schistosoma species over the threshold, it was classified into the oldest temporal group. If a second gene had a homolog in the pufferfish, medaka, trout or zebrafish species but not in the first clade was placed in the second temporal group. The species used for each temporal group are: TG1 (African malaria mosquito, fruitfly, nematode, Schistosoma and yellow fever mosquito), TG2 (medaka, pufferfish, trout and zebrafish), TG3 (clawed frog and tropical frog), TG4 (all chicken genes not found in the above species).

BS-PCR validation

One microgram of genomic DNA from each biological replicate was bisulfite-converted by the EZ DNA Methylation-Gold™ Kit. Primers were designed to amplify target regions of the bisulfite-converted DNA for validation of the MethylC-Seq results. Then we randomly selected 10–15 TA clones for each PCR product and sequenced by Sanger sequencing. All the primers were listed in Additional file 15.

Availability of supporting data

The sequencing data from this study have been submitted to the NCBI Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo>) under accession no. GSE56975.

Additional files

Additional file 1: Correlation analysis of the two biological replicates.

Overlap of mCs in CG, CHG and CHH contexts between the two biological replicates. mCs were classified as unique to biological replicate 1 (yellow), unique to replicate 2 (red) or shared by both replicates (orange). The number of overlapped mCs in each category was listed, as well as the percentage of mCs unique within each biological replicate. (TIFF 362 kb)

Additional file 2: Number of cytosines and methylcytosines detected in the chicken genome. (DOCX 16 kb)

Additional file 3: Bisulfite sequencing validation results on CG sites. (XLS 72 kb)

Additional file 4: Methylcytosine density in a 10-kb window of all chromosomes in chickens. (PDF 8163 kb)

Additional file 5: List of promoter and gene body hypo- and hyper-methylated genes. (XLSX 63 kb)

Additional file 6: GO enrichment of promoter and gene body hyper- or hypo-methylated genes. (A-B) GO enrichment of hyper-methylated genes (methylation level $\geq 70\%$) and hypo-methylated genes (methylation level $\leq 30\%$) in promoter (A) and gene body (B). Annotations are grouped by biological process, cellular component and molecular function based on the Gene Ontology database (<http://www.geneontology.org/>). Gene numbers are listed for each category. (TIFF 657 kb)

Additional file 7: DAVID functional enrichment of promoter and gene body hypo- and hyper-methylated genes. (XLSX 80 kb)

Additional file 8: The relative methylation level of different repeat types and genome random selected regions. (DOC 43 kb)

Additional file 9: DMR-associated genes between Fayoumi and Leghorn lines. (XLSX 205 kb)

Additional file 10: Expression levels of DMR-associated genes from RNA-seq data. (XLSX 161 kb)

Additional file 11: DMR-associated genes that show more than two-fold expression differences between Fayoumi and Leghorn. (XLSX 33 kb)

Additional file 12: Methylation distribution of a specific DMR-associated gene. (A) Methylation distribution of DMR-associated gene *HCK* (ENSGALG00000006522) was performed by UCSC genome browser custom tracks. Fayoumi and Leghorn lines were demonstrated separately. The bars indicate the methylation level of each mCG sites. Differentially methylated region was highlighted in light gray. (B) Bisulphite-PCR validation of some different methylated CG sites within the light gray region. Orange part represents the methylation percentage of each site. (C) RNA-seq result of *HCK* in Fayoumi and Leghorn birds. (TIFF 1128 kb)

Additional file 13: Methylation distribution and expression level of DMR-associated gene *TLR4* and *PIK3CD* in Fayoumi and Leghorn. (A) Methylation distribution of the differentially methylated regions in *TLR4* and *PIK3CD* by the UCSC genome browser custom track. Fayoumi and Leghorn lines were demonstrated separately. The bars indicate the methylation level of each mCG sites. Differentially methylated regions were highlighted in light gray. (B) RNA-seq result of the DMR-associated gene between Fayoumi and Leghorn lines. (PDF 1737 kb)

Additional file 14: Correlation analysis between DNA methylation differences and gene expression differences. X axis represented expression fold change and Y axis represented promoter methylation fold change of each gene. (PDF 95 kb)

Additional file 15: PCR primers used for Bisulfite sequencing validation. (XLSX 12 kb)

Abbreviations

DMR: Differentially methylated region; TSS: Transcriptional start site; TTS: Transcriptional terminal site; TG: Temporal group.

Competing interests

The authors declare that no competing interests exist.

Authors' contributions

Conceived and designed the experiments: SH, XH, HZ, and NL. Performed the experiments: JL, QL, YW, and LL. Analyzed the data: JL, RL, YW, FL. Contributed reagents/materials/analysis tools: RL, YW, FL, SL, and HZ. Wrote the paper: JL, RL, FL, HZ, NL. Provided comments on the manuscript: YZ, XG, CF, LL, SL, HZ, and NL. All authors read and approved the final manuscript.

Acknowledgments

This work was financially supported by the National High Technology Research and Development Program of China (No. 2013AA102501), National Natural Science Foundation of China (31000584) and the USDA National Institute of Food and Agriculture, Multi-State/Zhou project1002113 and Hatch project #5357 (SL).

Author details

¹The State Key Laboratory for Agro-biotechnology, China Agricultural University, Beijing 100193, China. ²Core Genomic Facility, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China. ³Department of Animal Science, University of California, Davis, CA 95616, USA. ⁴Department of Animal Science, Iowa State University, Ames, IA 50011, USA. ⁵CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China. ⁶Department of Poultry Science, Texas A&M University, College Station, TX 77845, USA. ⁷National Engineering Laboratory for Animal Breeding, China Agricultural University, Beijing 100193, China. ⁸College of Animal Science and Technology, Yunnan Agricultural University, Kunming, Yunnan, China.

Received: 13 October 2014 Accepted: 15 October 2015

Published online: 23 October 2015

References

- Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev.* 2002;16(1):6–21.
- Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, et al. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet.* 2007;39(4):457–66.
- Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature.* 2009;462(7271):315–22.
- Harris RA, Wang T, Coarfa C, Nagarajan RP, Hong C, Downey SL, et al. Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nat Biotechnol.* 2010;28(10):1097–105.
- Laird PW. Principles and challenges of genomewide DNA methylation analysis. *Nat Rev Genet.* 2010;11(3):191–203.
- Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, et al. Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell.* 2008;133(3):523–36.
- Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, et al. Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature.* 2008;452(7184):215–19.
- Xiang H, Zhu J, Chen Q, Dai F, Li X, Li M, et al. Single base-resolution methylome of the silkworm reveals a sparse epigenomic map. *Nat Biotechnol.* 2010;28(5):516–20.
- Mugal CF, Arndt PF, Holm L, Ellegren H. Evolutionary consequences of DNA methylation on the GC content in vertebrate genomes. *G3.* 2015;5(3):441–7.
- Burt DW. Emergence of the chicken as a model organism: implications for agriculture and biology. *Poult Sci.* 2007;86(7):1460–71.
- International Chicken Genome Sequencing C. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature.* 2004;432(7018):695–716.
- Burt DW, Bruley C, Dunn IC, Jones CT, Ramage A, Law AS, et al. The dynamics of chromosome evolution in birds and mammals. *Nature.* 1999;402(6760):411–3.
- Burt DW. Chicken genome: current status and future opportunities. *Genome Res.* 2005;15(12):1692–8.
- Huang Y, Li Y, Burt DW, Chen H, Zhang Y, Qian W, et al. The duck genome and transcriptome provide insight into an avian influenza virus reservoir species. *Nat Genet.* 2013;45(7):776–83.
- Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet.* 2003;33(Suppl):245–54.

16. Litt MD, Simpson M, Gaszner M, Allis CD, Felsenfeld G. Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science*. 2001;293(5539):2453–5.
17. Rincon-Arango H, Guerrero G, Valdes-Quezada C, Recillas-Targa F. Chicken alpha-globin switching depends on autonomous silencing of the embryonic pi globin gene by epigenetics mechanisms. *Journal of cellular biochemistry* 2009, 108(3):675–87.
18. Li Q, Li N, Hu X, Li J, Du Z, Chen L, et al. Genome-wide mapping of DNA methylation in chicken. *PLoS One*. 2011;6(5):e19428.
19. Tian F, Zhan F, Vanderkraats ND, Hiken JF, Edwards JR, Zhang H, et al. DNMT gene expression and methylome in Marek's disease resistant and susceptible chickens prior to and following infection by MDV. *Epigenetics: official journal of the DNA Methylation Society*. 2013;8(4):431–44.
20. Hu Y, Xu H, Li Z, Zheng X, Jia X, Nie Q, et al. Comparison of the genome-wide DNA methylation profiles between fast-growing and slow-growing broilers. *PLoS One*. 2013;8(2):e56411.
21. Lakshmanan N, Gavora JS, Lamont SJ. Major histocompatibility complex class II DNA polymorphisms in chicken strains selected for Marek's disease resistance and egg production or for egg production alone. *Poult Sci*. 1997;76(11):1517–23.
22. Cheeseman JH, Kaiser MG, Ciraci C, Kaiser P, Lamont SJ. Breed effect on early cytokine mRNA expression in spleen and cecum of chickens with and without *Salmonella enteritidis* infection. *Dev Comp Immunol*. 2007;31(1):52–60.
23. Zhou H, Lamont SJ. Global gene expression profile after *Salmonella enterica* Serovar enteritidis challenge in two F8 advanced intercross chicken lines. *Cytogenet Genome Res*. 2007;117(1–4):131–8.
24. Abasht B, Kaiser MG, van der Poel J, Lamont SJ. Genetic lines differ in Toll-like receptor gene expression in spleens of chicks inoculated with *Salmonella enterica* serovar Enteritidis. *Poult Sci*. 2009;88(4):744–9.
25. Wang Y, Lupiani B, Reddy SM, Lamont SJ, Zhou H. RNA-seq analysis revealed novel genes and signaling pathway associated with disease resistance to avian influenza virus infection in chickens. *Poult Sci*. 2014;93(2):485–93.
26. Zhou H, Lamont SJ. Genetic characterization of biodiversity in highly inbred chicken lines by microsatellite markers. *Anim Genet*. 1999;30(4):256–64.
27. Ferrari R, Berk AJ, Kurdistani SK. Viral manipulation of the host epigenome for oncogenic transformation. *Nat Rev Genet*. 2009;10(5):290–4.
28. Wu SC, Zhang Y. Active DNA demethylation: many roads lead to Rome. *Nat Rev Mol Cell Biol*. 2010;11(9):607–20.
29. Lamont SJ, Chen Y, Aarts HJ, van der Hulst-van Arkel MC, Beuving G, Leenstra FR. Endogenous viral genes in thirteen highly inbred chicken lines and in lines selected for immune response traits. *Poult Sci*. 1992;71(3):530–8.
30. Cheeseman JH, Kaiser MG, Lamont SJ. Genetic line effect on peripheral blood leukocyte cell surface marker expression in chickens. *Poult Sci*. 2004;83(6):911–6.
31. Wang Y, Lupiani B, Reddy SM, Lamont SJ, Zhou H. RNA-seq analysis revealed novel genes and signaling pathway associated with disease resistance to avian influenza virus infection in chickens. *Poult Sci*. 2013;93(2):485.
32. Bock C, Paulsen M, Tierling S, Mikeska T, Lengauer T, Walter J. CpG island methylation in human lymphocytes is highly correlated with DNA sequence, repeats, and predicted DNA structure. *PLoS Genet*. 2006;2(3):e26.
33. Straussman P, Nejman D, Roberts D, Steinfeld I, Blum B, Benvenisty N, et al. Developmental programming of CpG island methylation profiles in the human genome. *Nat Struct Mol Biol*. 2009;16(5):564–71.
34. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012;13(7):484–92.
35. Fazzari MJ, Greally JM. Epigenomics: beyond CpG islands. *Nat Rev Genet*. 2004;5(6):446–55.
36. He X, Chang S, Zhang J, Zhao Q, Xiang H, Kusonmano K, et al. MethyCancer: the database of human DNA methylation and cancer. *Nucleic Acids Res*. 2008;36(Database issue):D836–41.
37. Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, et al. Genome-wide high-resolution mapping and functional analysis of DNA methylation in Arabidopsis. *Cell*. 2006;126(6):1189–201.
38. Zhao Y, Mooney SD. Functional organization and its implication in evolution of the human protein-protein interaction network. *BMC Genomics*. 2012;13:150.
39. Zheng D, Frankish A, Baertsch R, Kapranov P, Reymond A, Choo SW, et al. Pseudogenes in the ENCODE regions: consensus annotation, analysis of transcription, and evolution. *Genome Res*. 2007;17(6):839–51.
40. Balasubramanian S, Zheng D, Liu YJ, Fang G, Frankish A, Carriero N, et al. Comparative analysis of processed ribosomal protein pseudogenes in four mammalian genomes. *Genome Biol*. 2009;10(1):R2.
41. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet*. 2008;9(6):465–76.
42. Zemach A, McDaniel IE, Silva P, Zilberman D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science*. 2010;328(5980):916–9.
43. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet*. 2010;11(3):204–20.
44. Holliday R, Pugh JE. DNA modification mechanisms and gene activity during development. *Science*. 1975;187(4173):226–32.
45. Wang X, Wheeler D, Avery A, Rago A, Choi JH, Colbourne JK, et al. Function and Evolution of DNA Methylation in *Nasonia vitripennis*. *PLoS Genet*. 2013;9(10):e1003872.
46. Jjingo D, Conley AB, Yi SV, Lunyak VV, Jordan IK. On the presence and role of human gene-body DNA methylation. *Oncotarget*. 2012;3(4):462–74.
47. Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet*. 2007;39(1):61–9.
48. Smit AF. Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr Opin Genet Dev*. 1999;9(6):657–63.
49. Garrick D, Fiering S, Martin DI, Whitelaw E. Repeat-induced gene silencing in mammals. *Nat Genet*. 1998;18(1):56–9.
50. Walsh CP, Chaillet JR, Bestor TH. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat Genet*. 1998;20(2):116–7.
51. Fedoroff NV. Presidential address. Transposable elements, epigenetics, and genome evolution. *Science*. 2012;338(6108):758–67.
52. Feng S, Cokus SJ, Zhang X, Chen PY, Bostick M, Goll MG, et al. Conservation and divergence of methylation patterning in plants and animals. *Proc Natl Acad Sci U S A*. 2010;107(19):8689–94.
53. Alba MM, Castresana J. Inverse relationship between evolutionary rate and age of mammalian genes. *Mol Biol Evol*. 2005;22(3):598–606.
54. Wolf YI, Novichkov PS, Karev GP, Koonin EV, Lipman DJ. The universal distribution of evolutionary rates of genes and distinct characteristics of eukaryotic genes of different apparent ages. *Proc Natl Acad Sci U S A*. 2009;106(18):7273–80.
55. Takuno S, Gaut BS. Gene body methylation is conserved between plant orthologs and is of evolutionary consequence. *Proc Natl Acad Sci U S A*. 2013;110(5):1797–802.
56. Takuno S, Gaut BS. Body-methylated genes in Arabidopsis thaliana are functionally important and evolve slowly. *Mol Biol Evol*. 2012;29(1):219–27.
57. Soond DR, Bjorgo E, Moltu K, Dale VQ, Patton DT, Torgersen KM, et al. PI3K p110delta regulates T-cell cytokine production during primary and secondary immune responses in mice and humans. *Blood*. 2010;115(11):2203–13.
58. Okkenhaug K, Bilancio A, Farjot G, Priddle H, Sancho S, Peskett E, et al. Impaired B and T cell antigen receptor signaling in p110delta PI 3-kinase mutant mice. *Science*. 2002;297(5583):1031–4.
59. Gou Z, Liu R, Zhao G, Zheng M, Li P, Wang H, et al. Epigenetic modification of TLRs in leukocytes is associated with increased susceptibility to *Salmonella enteritidis* in chickens. *PLoS One*. 2012;7(3):e33627.
60. Li X, Zhu J, Hu F, Ge S, Ye M, Xiang H, et al. Single-base resolution maps of cultivated and wild rice methylomes and regulatory roles of DNA methylation in plant gene expression. *BMC Genomics*. 2012;13:300.
61. Vaughn MW, Tanurdzic M, Lippman Z, Jiang H, Carrasquillo R, Rabinowicz PD, et al. Epigenetic natural variation in Arabidopsis thaliana. *PLoS Biol*. 2007;5(7):e174.
62. Liang F, Tang B, Wang Y, Wang J, Yu C, Chen X, et al. WBSA: web service for bisulfite sequencing data analysis. *PLoS One*. 2014;9(1):e86707.
63. Dawid IB. 5-methylcytidylic acid: absence from mitochondrial DNA of frogs and HeLa cells. *Science*. 1974;184(4132):80–1.
64. Pollack Y, Kasir J, Shemer R, Metzger S, Szyf M. Methylation pattern of mouse mitochondrial DNA. *Nucleic Acids Res*. 1984;12(12):4811–24.
65. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol*. 2010;11(10):R106.
66. Huang D, Chang TR, Aggarwal A, Lee RC, Ehrlich HP. Mechanisms and dynamics of mechanical strengthening in ligament-equivalent fibroblast-populated collagen matrices. *Ann Biomed Eng*. 1993;21(3):289–305.
67. da Huang W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):13.
68. Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol*. 2003;4(5):3.